

信州大学
大学院総合理工学研究科

修士論文

離散無記憶通信路に対する
コストあたり通信路容量の算出

指導教員 西新 幹彦 准教授

学科 電子情報システム工学科
専攻 工学専攻
分野 電子情報システム学分野
学籍番号 18W2041F
氏名 川合 康太

2020 年 2 月 25 日

目次

1	はじめに	1
2	コスト付き通信路符号化問題	1
2.1	コスト制約付き通信路容量	2
2.2	コストあたり通信路容量	4
3	コストあたり通信路容量の数値的な計算に向けて	5
3.1	有本-Blahut アルゴリズム	5
3.2	コストあたり通信路容量を数値的に計算するアルゴリズム	7
4	最適性の証明	8
4.1	$\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi)$ を達成する φ	8
4.2	$\sup_{\mathbf{p}} I(\mathbf{p}; \varphi^{(N)})$ を達成する $\mathbf{p}$	9
5	収束性の証明	11
5.1	$I(\mathbf{p}^{(N+1)}; \varphi^{(N)})$ の収束性	11
5.2	$I(\mathbf{p}^{(N)}; \varphi^{(N)})$ の収束性	16
6	まとめ	17
	謝辞	17
	参考文献	17
付録 A	定理 1 の証明	19
A.1	順定理の証明	19
A.2	逆定理の証明	26
付録 B	定理 2 の証明	29
B.1	順定理の証明	29
B.2	逆定理の証明	31
B.3	式 (136) の証明	33

1 はじめに

通信システムとは、符号シンボルを送受信することによって情報を伝えるシステムである。一般に、符号シンボルの送信には何らかのコストがかかる。具体的には、符号シンボルとして、電圧や電磁波、磁気を用いる場合、いずれも電力が必要になる。また、1つのシンボルを送信するのにかかる時間もコストとみなされうる。

通信システムによって、各シンボルのコストは同一の場合もあるが、一般には、シンボルごとに異なる。例えば、AWGN 通信路では、入力シンボルは任意の実数値をとるが、多くの場合、その2乗は入力シンボルのコストとみなされる [1]。パケット間隔を用いた通信では、送信間隔はコストとみなされる [2]。

情報理論において、符号シンボルにコストが定義されている通信路符号化問題をコスト付き通信路符号化問題という。通常の通信路符号化では、次元として、符号シンボルあたりの情報量が問題となる通信路容量と呼ばれる概念が定義されているが、コスト付き通信路符号化では、通信路容量の次元として、同じく符号シンボルあたりの情報量が問題となるコスト制約付き通信路容量と、コストあたりの情報量が問題となるコストあたり通信路容量、と呼ばれる概念が定義されている。

これに関して、Gallager[3] は、コスト制約付き通信路容量を制約付き最適化問題の最適解として明らかにし、Verdú[1] は、この結果を用いて、コストあたり通信路容量を制約付き最適化問題の最適解として明らかにしている。

コストあたり通信路容量の数値的な計算については、従来研究を組み合わせることで一応計算可能ではある [4]。

一方で、通信路容量の最適化の形の類似性により、通信路容量を数値的に計算するアルゴリズムである有本-Blahut アルゴリズム [5][6][7] を多少変更すれば、より効率的な計算方法が存在することが期待できる。

本論文では、定常無記憶通信路に対して、有本-Blahut アルゴリズムに基づく、コストあたり通信路容量を数値的に計算するアルゴリズムを提案する。

なお、本論文で登場する指数・対数の底はすべてネイピア数である。

2 コスト付き通信路符号化問題

本研究に至る主要な背景として、コスト制約付き通信路容量およびコストあたり通信路容量に関する定理を以下に述べる。

2.1 コスト制約付き通信路容量

$\mathcal{X}$ を入力アルファベット, $\mathcal{Y}$ を出力アルファベットとして, 通信路の遷移確率と呼ぶ条件付き確率の組 $W = \{W_{yx}\}_{(x,y) \in \mathcal{X} \times \mathcal{Y}}$ を条件

$$\sum_y W_{yx} = 1, \quad x \in \mathcal{X} \quad (1)$$

を満たすように任意に定める. 入出力アルファベット $\mathcal{X}, \mathcal{Y}$ はともに有限であるとする.

入力系列

$$\mathbf{x} = x_1 x_2 \cdots x_n \in \mathcal{X}^n \quad (2)$$

が与えられたとき, 出力系列

$$\mathbf{y} = y_1 y_2 \cdots y_n \in \mathcal{Y}^n \quad (3)$$

の出現する条件付き確率 W_{yx}^n が

$$W_{yx}^n = \prod_{i=1}^n W_{y_i x_i} \quad (4)$$

で与えられる通信路を定常無記憶通信路 $\mathbf{W}$ と呼び, $\mathbf{W} = \{W\}$ と表す. このような通信路を通した情報のやりとりは次のようにして行われる. すなわち, 伝送すべきメッセージの集合を $\mathcal{M}_n = \{1, 2, \dots, M_n\}$ としたとき, それらを長さ n の通信路入力に変換する写像 $\varphi_n: \mathcal{M}_n \rightarrow \mathcal{X}^n$ を符号器と呼ぶ. ここで, $\mathbf{u}_i = \varphi_n(i)$ をメッセージ i の符号器と呼び, $\mathcal{C}_n = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{M_n}\}$ を符号と呼ぶ. 符号語 φ_n はメッセージ i を送りたいとき, その符号語 $\mathbf{u}_i$ を通信路に入力する. 一方, 通信路の受信側では, あらかじめ定めておいた $\mathcal{Y}^n$ の排反な分割

$$\mathcal{Y}^n = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \cdots \cup \mathcal{D}_{M_n}, \quad \mathcal{D}_i \cap \mathcal{D}_j = \emptyset, \quad i \neq j \quad (5)$$

に基づいて, 受けとった出力 $\mathbf{y}$ が $\mathbf{y} \in \mathcal{D}_i$ ならば $i \in \mathcal{M}_n$ が送信されてきたものと判定する ($\mathcal{D}_i$ をメッセージ i の復号領域と呼ぶ). このような操作を復号といい, これを表す写像 $\psi_n: \mathcal{Y}^n \rightarrow \mathcal{M}_n$ を復号器という. また, 通信路符号化における復号誤り確率 ε_n を

$$\varepsilon_n \triangleq \frac{1}{M_n} \sum_{i \in \mathcal{M}_n} W_{\mathcal{D}_i^c}^n \mathbf{u}_i \quad (6)$$

と定める. ここで, ε_n はメッセージ $i \in \mathcal{M}_n$ が等確率で発生すると仮定したときの平均誤り確率であることに注意する. メッセージの集合 $\mathcal{M}_n$ のサイズが M_n , 誤り確率が ε_n である符

号 $\mathcal{C}_n = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{M_n}\}$ は (n, M_n, ε_n) 符号と呼ばれる．通信路符号化は通常，誤り確率をある与えられた基準値以下にするような符号器と復号器の組 (φ_n, ψ_n) が存在するという条件のもとで符号化レートを最大限どこまで大きくできるかという形に定式化される．まず最初に考えられるのは，誤り確率が $\varepsilon_n \rightarrow 0 (n \rightarrow \infty)$ となることを要求した場合である．

通信路符号化問題のある状況では，符号語のコストを考えなければならないことがある．一般に， $n = 1, 2, \dots$ に対して，写像 $c_n : \mathcal{X}^n \rightarrow \mathbb{R}$ を任意に定め，

$$\mathbf{c} = \{c_n\}_{n=1}^{\infty} \quad (7)$$

とおく．ただし， $\mathbb{R}$ は実数全体の集合を表す．このとき， $\mathbf{x} \in \mathcal{X}^n$ に対して $c_n(\mathbf{x})$ を $\mathbf{x}$ のコストと呼ぶ．本論文では，ある $c : \mathcal{X} \rightarrow \mathbb{R}$ を用いて

$$c_n(\mathbf{x}) = \sum_{i=1}^n c(x_i), \quad \mathbf{x} = x_1 x_2 \cdots x_n \quad (8)$$

と書ける特別な場合を考える．このようなコストは加法的コストと呼ばれる [8]．

Gallager[3] は 1968 年にコスト制約付き通信路容量に関して以下の定理を明らかにしている．

定義 1 レート R が Γ -達成可能とは，ある固定長符号の列 $\{\varphi_n\}_{n=1}^{\infty}$ が存在して

$$\frac{1}{n} c_n(\varphi_n(i)) \leq \Gamma, \quad i = 1, 2, \dots, M_n \quad (9)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \geq R \quad (10)$$

$$\lim_{n \rightarrow \infty} \varepsilon_n = 0 \quad (11)$$

を満たすことである．

定義 2 コスト制約付き通信路容量を

$$C(\Gamma) \triangleq \sup\{R \mid \text{レート } R \text{ が } \Gamma\text{-達成可能}\} \quad (12)$$

と定義する．

ここで，入力分布 $p = \{p_x\}_{x \in \mathcal{X}}$ を

$$\sum_x p_x = 1 \quad (13)$$

を満たすように任意に定める． $q = \{q_y\}_{y \in \mathcal{Y}}$ は $q_y = \sum_x p_x W_{yx}$ を満たし， p を入力分布としたときの通信路 W の出力分布を表す．

確率変数 X, Y に対して $I(X; Y)$ を

$$I(X; Y) \triangleq \sum_{x, y} p_x W_{yx} \log \frac{W_{yx}}{q_y} \quad (14)$$

と定義し、 X と Y の間の相互情報量と呼ぶ。

すると、次の定理が成り立つ。

定理 1 (Gallager [3]) 定常無記憶通信路 $\mathbf{W} = \{W\}$ のコスト制約付き通信路容量 $C(\Gamma)$ は

$$C(\Gamma) = \sup_{X: E[c(X)] \leq \Gamma} I(X; Y) \quad (15)$$

で与えられる。ここで、 Y は X を入力変数としたときの通信路 W の出力変数を表す。

2.2 コストあたり通信路容量

定常無記憶通信路 $\mathbf{W}$ を通した情報のやりとりは次のようにして行われる。すなわち、伝送すべきメッセージの集合を $\mathcal{M}_\nu = \{1, 2, \dots, M_\nu\}$ としたとき、それらを長さ n の通信路入力に変換する写像 $\varphi_\nu: \mathcal{M}_\nu \rightarrow \mathcal{X}^n$ を符号器と呼ぶ。ここで、 $\mathbf{u}_i = \varphi_\nu(i)$ をメッセージ i の符号語と呼び、 $\mathcal{C}_\nu = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{M_\nu}\}$ を符号と呼ぶ。符号器 φ_ν はメッセージ i を送りたいとき、その符号語 $\mathbf{u}_i$ を通信路に入力する。一方、通信路の受信側では、あらかじめ定めておいた $\mathcal{Y}^n$ の排反な分割

$$\mathcal{Y}^n = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \dots \cup \mathcal{D}_{M_\nu}, \quad \mathcal{D}_i \cap \mathcal{D}_j = \emptyset, \quad i \neq j \quad (16)$$

に基づいて、受けとった出力 $\mathbf{y}$ が $\mathbf{y} \in \mathcal{D}_i$ ならば $i \in \mathcal{M}_\nu$ が送信されてきたものと判定する ($\mathcal{D}_i$ をメッセージ i の復号領域と呼ぶ)。このような操作を復号といい、これを表す写像 $\psi_n: \mathcal{Y}^n \rightarrow \mathcal{M}_\nu$ を復号器という。また、通信路符号化における復号誤り確率 ε_ν を

$$\varepsilon_\nu \triangleq \frac{1}{M_\nu} \sum_{i \in \mathcal{M}_\nu} W_{\mathcal{D}_i^c}^n \mathbf{u}_i \quad (17)$$

と定める。ここで、 ε_ν はメッセージ $i \in \mathcal{M}_\nu$ が等確率で発生すると仮定したときの平均誤り確率であることに注意する。メッセージの集合 $\mathcal{M}_\nu$ のサイズが M_ν 、誤り確率が ε_ν である符号 $\mathcal{C}_\nu = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{M_\nu}\}$ は $(\nu, M_\nu, \varepsilon_\nu)$ 符号と呼ばれる。通信路符号化は通常、誤り確率をある与えられた基準値以下にするような符号器と復号器の組 (φ_ν, ψ_ν) が存在するという条件のもとで符号化レートを最大限どこまで大きくできるかという形に定式化される。まず最初に考えられるのは、誤り確率が $\varepsilon_\nu \rightarrow 0 (\nu \rightarrow \infty)$ となることを要求した場合である。

Verdú[1] は上記の Gallager[3] による結果を用いて、コストあたり通信路容量に関して以下の定理を明らかにしている。

定義 3 コストあたりレート R が達成可能とは、ある固定長符号の列 $\{\varphi_\nu\}_{\nu>0}$ が存在して

$$c_n(\varphi_\nu(i)) \leq \nu, \quad i = 1, 2, \dots, M_\nu \quad (18)$$

$$\liminf_{\nu \rightarrow \infty} \frac{1}{\nu} \log M_\nu \geq R \quad (19)$$

$$\lim_{\nu \rightarrow \infty} \varepsilon_\nu = 0 \quad (20)$$

を満たすことである。

定義 4 コストあたり通信路容量を

$$C \triangleq \sup\{R \mid \text{コストあたりレート } R \text{ が達成可能}\} \quad (21)$$

と定義する。

このとき、次の定理が成り立つ。

定理 2 (Verdú[1]) 定常無記憶通信路 $\mathbf{W} = \{W\}$ のコストあたり通信路容量 C は、

$$C = \sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} = \sup_X \frac{I(X; Y)}{E[c(X)]} \quad (22)$$

で与えられる。

ここで、 $\frac{I(X; Y)}{E[c(X)]}$ をコストあたり相互情報量と呼ぶ。

3 コストあたり通信路容量の数値的な計算に向けて

式 (22) からコストあたり通信路容量を数値的に計算するために、通信路容量やレート歪み関数を数値的に算出するアルゴリズムである有本-Blahut アルゴリズム [5][6][7] に基づき、コストあたり通信路容量を数値的に計算するアルゴリズムを提案する。このアルゴリズムを用いることで、コストの定義された任意の離散無記憶通信路に対して、コストあたり通信路容量に収束する値を求めることができる。

本章では、有本-Blahut アルゴリズムを紹介した後、これに基づいたコストあたり通信路容量を数値的に計算するアルゴリズムを提案する。

3.1 有本-Blahut アルゴリズム

通信路容量

$$C = \sup_X I(X; Y) \quad (23)$$

を数値的に計算するアルゴリズムである有本-Blahut アルゴリズムを以下に記す。

確率変数 X に対して $H(X)$ を

$$H(X) \triangleq - \sum_x p_x \log p_x \quad (24)$$

と定義し, X のエントロピーと呼ぶ. $\varphi = \{\varphi_{xy}\}_{(x,y) \in \mathcal{X} \times \mathcal{Y}}$ を条件

$$\sum_x \varphi_{xy} = 1, \quad y \in \mathcal{Y} \quad (25)$$

を満たすように任意に定める.

相互情報量は

$$I(X; Y) = H(X) + \sum_x p_x \sum_y W_{yx} \log \frac{p_x W_{yx}}{q_y} \quad (26)$$

とも表されることから, 相互情報量を拡張した概念として, 次の関数

$$I(\mathbf{p}; \boldsymbol{\varphi}) = H(X) + \sum_x p_x \sum_y W_{yx} \log \varphi_{xy} \quad (27)$$

を定義する.

1. 初期入力分布 $\mathbf{p}^{(1)} = \{p_x^{(1)}\}$ を内点にとる. 一様分布

$$p_x^{(1)} = \frac{1}{|\mathcal{X}|} \quad (28)$$

としてよい.

2. $N = 1$ とする.

3. $\sup_{\boldsymbol{\varphi}} I(\mathbf{p}^{(N)}; \boldsymbol{\varphi})$ を達成する $\boldsymbol{\varphi}$ を求める. 具体的には, これを $\boldsymbol{\varphi}^{(N)} = \{\varphi_{xy}^{(N)}\}$ とし,

$$\varphi_{xy}^{(N)} = \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \quad (29)$$

とする.

4. $\sup_{\mathbf{p}} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ を達成する $\mathbf{p}$ を求める. 具体的には, これを $\mathbf{p}^{(N+1)} = \{p_x^{(N+1)}\}$ とし,

$$p_x^{(N+1)} = \frac{\alpha_x^{(N)}}{\sum_x \alpha_x^{(N)}} \quad (30)$$

とする. ただし,

$$\alpha_x^{(N)} = \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} \right) \quad (31)$$

とする.

5. N を 1 増やして 3. に戻る.

以上より得られる

$$\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi) = I(\mathbf{p}^{(N)}; \varphi^{(N)}) \quad (32)$$

および

$$\sup_{\mathbf{p}} I(\mathbf{p}; \varphi^{(N)}) = I(\mathbf{p}^{(N+1)}; \varphi^{(N)}) \quad (33)$$

は通信路容量 C に収束する値となる。

したがって、ループの回数 N を十分大きくすることによって誤差の少ない値が得られる。誤差 $C - I(\mathbf{p}^{(N+1)}; \varphi^{(N)})$ は高々 $1/N$ のオーダーであり、精度として小数第 $\log_{10} N - 1$ 桁まで保証される。

3.2 コストあたり通信路容量を数値的に計算するアルゴリズム

コストあたり通信路容量

$$C = \sup_X \frac{I(X; Y)}{E[c(X)]} \quad (34)$$

を数値的に計算するアルゴリズムを以下に提案する。

3.1 節と同様に、コストあたり相互情報量を拡張した概念として、次の関数

$$I(\mathbf{p}; \varphi) = \frac{1}{E[c(X)]} \left(H(X) + \sum_x p_x \sum_y W_{yx} \log \varphi_{xy} \right) \quad (35)$$

を定義する。

1. 初期入力分布 $\mathbf{p}^{(1)} = \{p_x^{(1)}\}$ を内点にとる。一様分布

$$p_x^{(1)} = \frac{1}{|\mathcal{X}|} \quad (36)$$

としてよい。

2. $N = 1$ とする。

3. $\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi)$ を達成する φ を求める。具体的には、これを $\varphi^{(N)} = \{\varphi_{xy}^{(N)}\}$ とし、

$$\varphi_{xy}^{(N)} = \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \quad (37)$$

とする。

4. $\sup_{\mathbf{p}} I(\mathbf{p}; \varphi^{(N)})$ を達成する $\mathbf{p}$ を求める。具体的には、これを $\mathbf{p}^{(N+1)} = \{p_x^{(N+1)}\}$ とし、

$$p_x^{(N+1)} = \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \alpha^{(N)} c(x) \right) \quad (38)$$

とする。ただし,

$$\alpha^{(N)} : \sum_x \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \alpha^{(N)} c(x) \right) = 1 \quad (39)$$

とする。

5. N を 1 増やして 3. に戻る。

以上より得られる

$$\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi) = I(\mathbf{p}^{(N)}; \varphi^{(N)}) \quad (40)$$

および

$$\sup_{\mathbf{p}} I(\mathbf{p}; \varphi^{(N)}) = I(\mathbf{p}^{(N+1)}; \varphi^{(N)}) \quad (41)$$

はコストあたり通信路容量 C に収束する値となる。

したがって、ループの回数 N を十分大きくすることによって誤差の少ない値が得られる。
ループを有限回で打ち切ったときの精度については現時点ではよく分かっていない。

4 最適性の証明

本章では、 $\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi)$ を達成する φ および $\sup_{\mathbf{p}} I(\mathbf{p}; \varphi^{(N)})$ を達成する $\mathbf{p}$ がそれぞれ式 (37), (38) を満たすことを証明する。

4.1 $\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi)$ を達成する φ

定理 3 $\sup_{\varphi} I(\mathbf{p}^{(N)}; \varphi)$ を達成する φ は,

$$\varphi_{xy}^{(N)} = \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \quad (42)$$

となる。

証明 カルバック・ライブラー情報量の性質より,

$$\begin{aligned}
I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}) &= \frac{1}{E[c(X^{(N)})]} \left(H(X^{(N)}) + \sum_x p_x^{(N)} \sum_y W_{yx} \log \varphi_{xy} \right) \\
&= \frac{1}{E[c(X^{(N)})]} \left(H(X^{(N)}) + \sum_y q_y^{(N)} \sum_x \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \log \varphi_{xy} \right) \\
&\leq \frac{1}{E[c(X^{(N)})]} \left(H(X^{(N)}) + \sum_y q_y^{(N)} \sum_x \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \log \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \right) \\
&= \frac{1}{E[c(X^{(N)})]} \left(H(X^{(N)}) + \sum_x p_x^{(N)} \sum_y W_{yx} \log \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \right) \\
&= \frac{I(X^{(N)}; Y^{(N)})}{E[c(X^{(N)})]} \tag{43}
\end{aligned}$$

を得る. 式 (43) は φ_{xy} に無関係な定数であるから, $\sup_{\boldsymbol{\varphi}} I(\mathbf{p}^{(N)}; \boldsymbol{\varphi})$ を達成する $\boldsymbol{\varphi}$ は

$$\varphi_{xy}^{(N)} = \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}} \tag{44}$$

となる. $\square$

4.2 $\sup_{\mathbf{p}} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ を達成する $\mathbf{p}$

定理 4 $\sup_{\mathbf{p}} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ を達成する $\mathbf{p}$ は,

$$p_x^{(N+1)} = \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \alpha^{(N)} c(x) \right) \tag{45}$$

となる. ただし, $\alpha^{(N)}$ は,

$$\sum_x \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \alpha^{(N)} c(x) \right) = 1 \tag{46}$$

を満たす唯一の実数解である.

証明 制約条件

$$p_x \geq 0, \quad x \in \mathcal{X} \tag{47}$$

$$\sum_x p_x = 1 \tag{48}$$

の下で，目的関数 $I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ の停留点を求める．ラグランジュ関数を

$$J(\mathbf{p}, \lambda) = I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) + \lambda \left(\sum_x p_x - 1 \right) \quad (49)$$

とにおいて，ラグランジュの未定乗数法を適用する．ただし， λ は，ラグランジュ乗数である． $J(\mathbf{p}, \lambda)$ を p_x で偏微分し，これを 0 とおくと

$$\frac{1}{E[c(X)]} \left(-\log p_x - 1 + \sum_y W_{yx} \log \varphi_{xy}^{(N)} \right) - \frac{1}{(E[c(X)])^2} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) c(x) + \lambda = 0 \quad (50)$$

となり，両辺に $E[c(X)]$ を掛けて

$$-\log p_x - 1 + \sum_y W_{yx} \log \varphi_{xy}^{(N)} - \frac{1}{E[c(X)]} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) c(x) + \lambda E[c(X)] = 0 \quad (51)$$

を得る．両辺に p_x を掛け， x で和をとると

$$H(X) - 1 + \sum_x p_x \sum_y W_{yx} \log \varphi_{xy}^{(N)} - I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) + \lambda E[c(X)] = 0 \quad (52)$$

となり，整理すると，

$$-1 + \lambda E[c(X)] = 0 \quad (53)$$

となり， λ について解くと

$$\lambda = \frac{1}{E[c(X)]} \quad (54)$$

となる．これを式 (51) に代入して

$$-\log p_x + \sum_y W_{yx} \log \varphi_{xy}^{(N)} - \frac{1}{E[c(X)]} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) c(x) = 0 \quad (55)$$

となり， p_x について解くと

$$p_x = \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \frac{1}{E[c(X)]} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) c(x) \right) \quad (56)$$

を得る．両辺を x で和をとると

$$1 = \sum_x \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \frac{1}{E[c(X)]} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)}) c(x) \right) \quad (57)$$

となる．ここで，式 (57) は，未知数 $I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ の方程式であることに注意して，その唯一の実数解を $\alpha^{(N)}$ とおけば，式 (56) より，入力分布は，

$$p_x^{(N+1)} = \exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \alpha^{(N)} c(x) \right) \quad (58)$$

となる．このとき，式 (58) が目的関数 $I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ の極大点であれば，これは $\sup_{\mathbf{p}} I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ を達成する入力分布となる．□

5 収束性の証明

本章では， $I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})$ および $I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)})$ がコストあたり通信路容量 C に収束することを証明する．

5.1 $I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})$ の収束性

次の補題を準備する．

補題 1 $\mathbf{p}^{(1)}$ は内点にあるとし，コストあたり相互情報量 $\frac{I(X;Y)}{E[c(X)]}$ を最大にする入力確率分布の 1 つを $\mathbf{p}^{(0)}$ とする．このとき，等式

$$\sum_x p_x^{(0)} \log \frac{p_x^{(N+1)}}{p_x^{(N)}} = \sum_y q_y^{(0)} \log \frac{q_y^{(0)}}{q_y^{(N)}} + E[c(X^{(0)})] \left(C - I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \right) \quad (59)$$

が成り立つ．

証明 定理 4 を利用して，補題 1 が次のように導かれる．式 (59) の左辺を変形して

$$\begin{aligned} \sum_x p_x^{(0)} \log \frac{p_x^{(N+1)}}{p_x^{(N)}} &= \sum_x p_x^{(0)} \log \left(\frac{p_x^{(N+1)}}{p_x^{(N)}} \frac{\exp(E[c(X^{(0)})]\alpha^{(N)})}{\exp(E[c(X^{(0)})]\alpha^{(N)})} \right) \\ &= \sum_x p_x^{(0)} \log \frac{1}{\exp(E[c(X^{(0)})]\alpha^{(N)})} \\ &\quad + \sum_x p_x^{(0)} \log \left(\frac{p_x^{(N+1)}}{p_x^{(N)}} \exp(E[c(X^{(0)})]\alpha^{(N)}) \right) \end{aligned} \quad (60)$$

を得る．ここで，式 (60) の第 1 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \frac{1}{\exp(E[c(X^{(0)})]\alpha^{(N)})} \\
&= - \sum_x p_x^{(0)} E[c(X^{(0)})]\alpha^{(N)} \\
&= - E[c(X^{(0)})]\alpha^{(N)} \\
&= - E[c(X^{(0)})]I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})
\end{aligned} \tag{61}$$

式 (60) の第 2 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \left(\frac{p_x^{(N+1)}}{p_x^{(N)}} \exp(E[c(X^{(0)})]\alpha^{(N)}) \right) \\
&= \sum_x p_x^{(0)} \log \left(\frac{\exp(\sum_y W_{yx} \log \varphi_{xy}^{(N)} - \alpha^{(N)} c(x))}{p_x^{(N)}} \exp(E[c(X^{(0)})]\alpha^{(N)}) \right) \\
&= \sum_x p_x^{(0)} \log \left(\frac{\exp(\sum_y W_{yx} \log \varphi_{xy}^{(N)})}{p_x^{(N)}} \exp(-\alpha^{(N)} (c(x) - E[c(X^{(0)})])) \right) \\
&= \sum_x p_x^{(0)} \log \frac{1}{p_x^{(N)}} \\
&\quad + \sum_x p_x^{(0)} \log \left(\exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} \right) \right) \\
&\quad + \sum_x p_x^{(0)} \log \left(\exp(-\alpha^{(N)} (c(x) - E[c(X^{(0)})])) \right)
\end{aligned} \tag{62}$$

となる．また，式 (62) の第 1 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \frac{1}{p_x^{(N)}} \\
&= \sum_x p_x^{(0)} \log \frac{1}{p_x^{(N)}} \sum_y W_{yx} \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{1}{p_x^{(N)}}
\end{aligned} \tag{63}$$

式 (62) の第 2 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \left(\exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} \right) \right) \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \varphi_{xy}^{(N)}
\end{aligned} \tag{64}$$

そして、式 (62) の第 3 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \left(\exp \left(-\alpha^{(N)} \left(c(x) - E[c(X^{(0)})] \right) \right) \right) \\
&= \sum_x p_x^{(0)} \left(-\alpha^{(N)} \left(c(x) - E[c(X^{(0)})] \right) \right) \\
&= -\alpha^{(N)} \sum_x p_x^{(0)} \left(c(x) - E[c(X^{(0)})] \right) \\
&= -\alpha^{(N)} \left(\sum_x c(x) p_x^{(0)} - E[c(X^{(0)})] \sum_x p_x^{(0)} \right) \\
&= -\alpha^{(N)} \left(E[c(X^{(0)})] - E[c(X^{(0)})] \right) \\
&= 0
\end{aligned} \tag{65}$$

となるから、式 (60) の第 2 項は、式 (63),(64),(65) を合わせて、

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \left(\frac{p_x^{(N+1)}}{p_x^{(N)}} \exp \left(E[c(X^{(0)})] \alpha^{(N)} \right) \right) \\
&= \sum_x p_x^{(0)} \log \frac{1}{p_x^{(N)}} \\
&\quad + \sum_x p_x^{(0)} \log \left(\exp \left(\sum_y W_{yx} \log \varphi_{xy}^{(N)} \right) \right) \\
&\quad + \sum_x p_x^{(0)} \log \left(\exp \left(-\alpha^{(N)} \left(c(x) - E[c(X^{(0)})] \right) \right) \right) \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{1}{p_x^{(N)}} \quad (\because (63)) \\
&\quad + \sum_x p_x^{(0)} \sum_y W_{yx} \log \varphi_{xy}^{(N)} \quad (\because (64)) \\
&\quad + 0 \quad (\because (65)) \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \left(\log \frac{1}{p_x^{(N)}} + \log \varphi_{xy}^{(N)} \right) \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{\varphi_{xy}^{(N)}}{p_x^{(N)}} \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \left(\frac{\varphi_{xy}^{(N)}}{p_x^{(N)}} \frac{q_y^{(0)}}{q_y^{(0)}} \frac{W_{yx}}{W_{yx}} \right) \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{W_{yx}}{q_y^{(0)}} + \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{q_y^{(0)} \varphi_{xy}^{(N)}}{p_x^{(N)} W_{yx}}
\end{aligned} \tag{66}$$

となる．さらに，式 (66) の第 1 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{W_{yx}}{q_y^{(0)}} \\
&= \frac{E[c(X^{(0)})]}{E[c(X^{(0)})]} \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{W_{yx}}{q_y^{(0)}} \\
&= E[c(X^{(0)})] \frac{I(X^{(0)}; Y^{(0)})}{E[c(X^{(0)})]} \\
&= E[c(X^{(0)})] \sup_X \frac{I(X; Y)}{E[c(X)]} \\
&= E[c(X^{(0)})] C
\end{aligned} \tag{67}$$

式 (66) の第 2 項は

$$\begin{aligned}
& \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{q_y^{(0)} \varphi_{xy}^{(N)}}{p_x^{(N)} W_{yx}} \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{q_y^{(0)} \frac{p_x^{(N)} W_{yx}}{q_y^{(N)}}}{p_x^{(N)} W_{yx}} \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{q_y^{(0)}}{q_y^{(N)}} \\
&= \sum_y \left(\sum_x p_x^{(0)} W_{yx} \right) \log \frac{q_y^{(0)}}{q_y^{(N)}} \\
&= \sum_y q_y^{(0)} \log \frac{q_y^{(0)}}{q_y^{(N)}}
\end{aligned} \tag{68}$$

となるから，式 (66) は，

$$\begin{aligned}
& \sum_x p_x^{(0)} \log \left(\frac{p_x^{(N+1)}}{p_x^{(N)}} \exp \left(E[c(X^{(0)})] \alpha^{(N)} \right) \right) \\
&= \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{W_{yx}}{q_y^{(0)}} + \sum_x p_x^{(0)} \sum_y W_{yx} \log \frac{q_y^{(0)} \varphi_{xy}^{(N)}}{p_x^{(N)} W_{yx}} \\
&= E[c(X^{(0)})] C + \sum_y q_y^{(0)} \log \frac{q_y^{(0)}}{q_y^{(N)}}
\end{aligned} \tag{69}$$

となる。したがって、式 (60) は、式 (61),(69) を合わせて、

$$\begin{aligned}
\sum_x p_x^{(0)} \log \frac{p_x^{(N+1)}}{p_x^{(N)}} &= \sum_x p_x^{(0)} \log \frac{1}{\exp(E[c(X^{(0)})]\alpha^{(N)})} \\
&\quad + \sum_x p_x^{(0)} \log \left(\frac{p_x^{(N+1)}}{p_x^{(N)}} \exp(E[c(X^{(0)})]\alpha^{(N)}) \right) \\
&= -E[c(X^{(0)})]I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \quad (\because (61)) \\
&\quad + E[c(X^{(0)})]C + \sum_y q_y^{(0)} \log \frac{q_y^{(0)}}{q_y^{(N)}} \quad (\because (69)) \\
&= \sum_y q_y^{(0)} \log \frac{q_y^{(0)}}{q_y^{(N)}} + E[c(X^{(0)})] \left(C - I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \right) \quad (70)
\end{aligned}$$

となる。□

この補題を利用して、 $I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})$ が C に収束することを証明する。

定理 5 $I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})$ は、コストあたり通信路容量 C に収束する。

証明 式 (59) の右辺の第 1 項は、カルバック・ライブラー情報量の性質より非負であるから、一般に不等式

$$E[c(X^{(0)})] \left(C - I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \right) \leq \sum_x p_x^{(0)} \log \frac{p_x^{(N+1)}}{p_x^{(N)}} \quad (71)$$

すなわち、

$$C - I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \leq \frac{1}{E[c(X^{(0)})]} \sum_x p_x^{(0)} \log \frac{p_x^{(N+1)}}{p_x^{(N)}}$$

が成立する。これを $N = 1$ から任意の自然数 K まで加えると、カルバック・ライブラー情報量の性質より、

$$\begin{aligned}
\sum_{N=1}^K \left(C - I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \right) &\leq \frac{1}{E[c(X^{(0)})]} \sum_x p_x^{(0)} \log \frac{p_x^{(K+1)}}{p_x^{(1)}} \\
&\leq \frac{1}{E[c(X^{(0)})]} \sum_x p_x^{(0)} \log \frac{p_x^{(0)}}{p_x^{(1)}} \quad (72)
\end{aligned}$$

が成立する．この右辺の項は， $\mathbf{p}^{(1)}$ が内点にあることから，有限な値になり，ここで，

$$\begin{aligned}
 I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) &\leq \sup_{\boldsymbol{\varphi}} I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}) \\
 &= I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N+1)}) \\
 &= \frac{I(X^{(N+1)}; Y^{(N+1)})}{E[c(X^{(N+1)})]} \\
 &\leq \sup_X \frac{I(X; Y)}{E[c(X)]} \\
 &= C
 \end{aligned} \tag{73}$$

となるから，式 (72) の左辺の和の中はすべて非負であり， $K \rightarrow \infty$ としてみればわかるように

$$\lim_{N \rightarrow \infty} I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) = C \tag{74}$$

でなければならない．□

5.2 $I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)})$ の収束性

5.1 節で証明した $I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})$ の収束性が保証されていれば十分であるが，本節では， $I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)})$ がコストあたり通信路容量 C に収束することを証明する．

定理 6 $I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)})$ は，コストあたり通信路容量 C に収束する．

証明

$$\begin{aligned}
 I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)}) &= \sup_{\boldsymbol{\varphi}} I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}) \\
 &\geq I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N-1)})
 \end{aligned} \tag{75}$$

また，

$$\begin{aligned}
 I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)}) &= \frac{I(X^{(N)}; Y^{(N)})}{E[c(X^{(N)})]} \\
 &\leq \sup_X \frac{I(X; Y)}{E[c(X)]} \\
 &= C
 \end{aligned} \tag{76}$$

となるから，

$$I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N-1)}) \leq I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)}) \leq C \tag{77}$$

が成り立つ．したがって，はさみうちの原理から，

$$\lim_{N \rightarrow \infty} I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)}) = C \tag{78}$$

となる．□

6 まとめ

本論文では、コストあたり通信路容量を数値的に計算するアルゴリズムを通信路容量を数値的に計算するアルゴリズムである有本-Blahut アルゴリズムに基づいて提案した。収束性は保証されているが、単調性の保証されていない部分が存在する。4.2 節で目的関数 $I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ の停留点 $\mathbf{p}^{(N+1)}$ が極大点といえていない部分である。よって、 $\mathbf{p}^{(N+1)}$ が鞍点または極小点となる可能性があり、アルゴリズムを計算している中で $I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)})$ の値が $I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)})$ の値より小さくなる可能性があり、次のような単調性が保証されない。

$$\begin{aligned} I(\mathbf{p}^{(1)}; \boldsymbol{\varphi}^{(1)}) &\leq I(\mathbf{p}^{(2)}; \boldsymbol{\varphi}^{(1)}) \leq I(\mathbf{p}^{(2)}; \boldsymbol{\varphi}^{(2)}) \leq \dots \\ &\leq I(\mathbf{p}^{(N)}; \boldsymbol{\varphi}^{(N)}) \leq I(\mathbf{p}^{(N+1)}; \boldsymbol{\varphi}^{(N)}) \leq \dots \leq C \end{aligned} \quad (79)$$

だが、目的関数 $I(\mathbf{p}; \boldsymbol{\varphi}^{(N)})$ は上に有界であり、停留点を 1 つしか持たないことから、おおかた $\mathbf{p}_x^{(N+1)}$ は極大点であると予想される。いずれにせよ、ループの回数 N が十分大きくなった時点の $\boldsymbol{\varphi}^{(N)}$ を用いて近似値を得ることができる。ループを有限回で打ち切ったときの精度については現時点ではよく分かっていない。このため、アルゴリズムの収束の速さを検討し、停止させるタイミングの目安を得ることが今後の課題である。

また、式 (22) の第 2 項より、ニュートン法を用いることで、コストあたり通信路容量は一応数値的には計算可能である [4]。この手法と提案アルゴリズムとの比較も今後の課題である。

謝辞

本研究を進めるにあたり、丁寧なご指導を頂いた西新幹彦准教授に感謝の意を表する。また、情報理論とその応用シンポジウムにおいて本研究を発表した際に助言を頂いた東京工業大学植松友彦教授にも感謝の意を表する。

参考文献

- [1] Sergio Verdú, “On Channel Capacity per Unit Cost,” *IEEE Transactions on Information Theory*, vol.36, no.5, pp.1019–1030, Sep. 1990.
- [2] Sergio Verdú, “Bits Through Queues,” *IEEE Transactions on Information Theory*, vol.42, no.1, pp.4–18, Jan. 1996.
- [3] Robert G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, 1968.

- [4] 三宅茂樹, 「電子情報通信学会 知識ベース 1 群 信号・システム 1 編 情報理論 8 章 通信路符号化」, http://www.ieice-hbkb.org/files/01/01gun_01hen_08.pdf#page=8, 2019 年 10 月閲覧.
- [5] Suguru Arimoto, “An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channels,” IEEE Transactions on Information Theory, vol.IT-18, no.1, pp.14–20, Jan. 1972.
- [6] Richard E. Blahut, “Computation of Channel Capacity and Rate-Distortion Functions,” IEEE Transactions on Information Theory, vol.IT-18, no.4, pp.460–473, Jul. 1972.
- [7] 有本卓, 情報理論, 共立出版, 1976.
- [8] 韓太舜, 情報理論における情報スペクトル的方法, 培風館, 1998.

付録 A 定理 1 の証明

定理 1 を順定理および逆定理に分けて証明する.

A.1 順定理の証明

次の補題を準備する.

補題 2

$$f(\Gamma) = \sup_{X: E[c(X)] \leq \Gamma} I(X; Y) \quad (80)$$

は上に凸である.

証明 任意の Γ_1, Γ_2 を考える.

$$E[c(X_i)] \leq \Gamma_i \quad (81)$$

なる X_1, X_2 を考える. また,

$$\alpha_1 + \alpha_2 = 1, \alpha_1 \geq 0, \alpha_2 \geq 0 \quad (82)$$

なる α_1, α_2 を考える.

$$P_X(x) = \alpha_1 P_{X_1}(x) + \alpha_2 P_{X_2}(x) \quad (83)$$

なる確率変数 X を考える. すると,

$$\begin{aligned} E[c(X)] &= \sum_x c(x) P_X(x) \\ &= \alpha_1 \sum_x c(x) P_{X_1}(x) + \alpha_2 \sum_x c(x) P_{X_2}(x) \\ &= \alpha_1 E[c(X_1)] + \alpha_2 E[c(X_2)] \\ &\leq \alpha_1 \Gamma_1 + \alpha_2 \Gamma_2. \end{aligned} \quad (84)$$

よって,

$$\begin{aligned}
\alpha_1 I(X_1; Y_1) + \alpha_2 I(X_2; Y_2) &= \sum_{i=1}^2 \alpha_i \sum_x P_{X_i}(x) \sum_y W(y|x) \log \frac{W(y|x)}{P_{Y_i}(y)} \\
&= \sum_x \sum_{i=1}^2 \alpha_i P_{X_i}(x) \sum_y W(y|x) \log W(y|x) \\
&\quad + \sum_{i=1}^2 \alpha_i \sum_y \sum_x P_{X_i}(x) W(y|x) \log \frac{1}{P_{Y_i}(y)} \\
&= \sum_x P_X(x) \sum_y W(y|x) \log W(y|x) \\
&\quad + \sum_{i=1}^2 \alpha_i \sum_y \sum_x P_{X_i Y_i}(x, y) \log \frac{1}{P_{Y_i}(y)} \\
&= \sum_{x, y} P_{XY}(x, y) \log W(y|x) \\
&\quad + \sum_{i=1}^2 \alpha_i \sum_y P_{Y_i}(y) \log \frac{1}{P_{Y_i}(y)} \\
&= -H(Y|X) + \sum_{i=1}^2 \alpha_i H(Y_i). \tag{85}
\end{aligned}$$

ここで,

$$P_Q(i) = \alpha_i \tag{86}$$

なる確率変数 Q を導入すると,

$$\begin{aligned}
P_Y(y) &= \sum_x P_X(x) W(y|x) \\
&= \sum_x \sum_{i=1}^2 \alpha_i P_{X_i}(x) W(y|x) \\
&= \sum_{i=1}^2 \alpha_i \sum_x P_{X_i}(x) W(y|x) \\
&= \sum_{i=1}^2 \alpha_i P_{Y_i}(y) \\
&= \sum_{i=1}^2 P_Q(i) P_{Y_i}(y) \tag{87}
\end{aligned}$$

であるので,

$$P_{Y_i}(y) = P_{Y|Q}(y|i) \tag{88}$$

と書くことができる。よって,

$$\begin{aligned} \sum_{i=1}^2 \alpha_i \sum_y P_{Y_i}(y) \log \frac{1}{P_{Y_i}(y)} &= \sum_{i=1}^2 P_Q(i) \sum_y P_{Y|Q}(y|i) \log \frac{1}{P_{Y|Q}(y|i)} \\ &= H(Y|Q) \end{aligned} \quad (89)$$

となるので, 式 (85) は,

$$\begin{aligned} \alpha_1 I(X_1; Y_1) + \alpha_2 I(X_2; Y_2) &= -H(Y|X) + H(Y|Q) \\ &= H(Y) - H(Y|X) + H(Y|Q) - H(Y) \\ &= I(X; Y) - I(Y; Q) \\ &\leq I(X; Y). \end{aligned} \quad (90)$$

よって,

$$\begin{aligned} \alpha_1 f(\Gamma_1) + \alpha_2 f(\Gamma_2) &= \sup_{X_1: E[c(X_1)] \leq \Gamma_1} \alpha_1 I(X_1; Y_1) + \sup_{X_2: E[c(X_2)] \leq \Gamma_2} \alpha_2 I(X_2; Y_2) \\ &= \sup_{X_1: E[c(X_1)] \leq \Gamma_1, X_2: E[c(X_2)] \leq \Gamma_2} (\alpha_1 I(X_1; Y_1) + \alpha_2 I(X_2; Y_2)) \\ &\leq \sup_{X_1: E[c(X_1)] \leq \Gamma_1, X_2: E[c(X_2)] \leq \Gamma_2} I(X; Y) \\ &\leq \sup_{X: E[c(X)] \leq \alpha_1 \Gamma_1 + \alpha_2 \Gamma_2} I(X; Y) \\ &= f(\alpha_1 \Gamma_1 + \alpha_2 \Gamma_2). \end{aligned} \quad (91)$$

よって, 上に凸である. $\square$

補題 2 を利用して, 順定理を証明する.

証明 任意の $\tilde{\gamma} > 0$ に対して,

$$E[c(X)] \leq \Gamma - \tilde{\gamma} \quad (92)$$

を満たす任意の確率変数 $X \in \mathcal{X}$ および Γ を考える. X_i ($i = 1, 2, \dots, n$) をそれぞれ X の i.i.d. コピーとし,

$$X^n = X_1 X_2 \cdots X_n \quad (93)$$

とおく. ここで,

$$\mathcal{A}_n = \left\{ \mathbf{x} \in \mathcal{X}^n \mid \frac{1}{n} c(\mathbf{x}) \leq \Gamma \right\} \quad (94)$$

$$\kappa_n = P_{X^n}(\mathcal{A}_n) \quad (95)$$

とすると,

$$\begin{aligned}
\kappa_n &= P_{X^n}(\mathcal{A}_n) \\
&= \Pr \left\{ \frac{1}{n} c(X^n) \leq \Gamma \right\} \\
&= \Pr \left\{ \frac{1}{n} \sum_{i=1}^n c(X_i) \leq \Gamma \right\} \\
&\geq \Pr \left\{ \frac{1}{n} \sum_{i=1}^n c(X_i) \leq E[c(X)] + \tilde{\gamma} \right\} \\
&= \Pr \left\{ \frac{1}{n} \sum_{i=1}^n c(X_i) - E[c(X)] \leq \tilde{\gamma} \right\} \\
&\geq \Pr \left\{ -\tilde{\gamma} \leq \frac{1}{n} \sum_{i=1}^n c(X_i) - E[c(X)] \leq \tilde{\gamma} \right\} \\
&= \Pr \left\{ \left| \frac{1}{n} \sum_{i=1}^n c(X_i) - E[c(X)] \right| \leq \tilde{\gamma} \right\} \\
&\rightarrow 1 \quad (n \rightarrow \infty) \quad (\because \text{大数の弱法則})
\end{aligned} \tag{96}$$

ここで,

$$P_{\tilde{X}^n}(\mathbf{x}) = \begin{cases} \frac{1}{\kappa_n} P_{X^n}(\mathbf{x}), & \mathbf{x} \in \mathcal{A}_n \\ 0, & \mathbf{x} \notin \mathcal{A}_n \end{cases} \tag{97}$$

を満たす確率変数 $\tilde{X}^n$ を定める. この $\tilde{X}^n$ で, 以下のように符号を作る. ここで, 任意の $\gamma > 0$ に対して,

$$\mathcal{T}_n = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \frac{1}{n} \log \frac{W^n(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} > I(X; Y) - 2\gamma \right\} \tag{98}$$

$$R = I(X; Y) - 3\gamma \tag{99}$$

$$M_n = e^{nR} \tag{100}$$

とおく. M_n 個の符号語を分布 $P_{\tilde{X}^n}$ に従って, 確率的に独立に決める (ランダム符号). これらの符号語を $\varphi_n(1), \varphi_n(2), \dots, \varphi_n(M_n)$ と表す. これで, 符号語 φ_n が定義することができた. 復号器は, $\mathbf{y} \in \mathcal{Y}^n$ を受信すると, $(\varphi_n(i), \mathbf{y}) \in \mathcal{T}_n$ を満たす $i \in \{1, 2, \dots, M_n\}$ が唯一存在するとき, i を復号メッセージとして出力する. そのような i が存在しないときや, 複数存在するときは, 復号誤りが起こったものとする. 復号誤り確率を評価するために,

$$E_i(\mathbf{y}) = \{(\varphi_n(i), \mathbf{y}) \in \mathcal{T}_n\} \tag{101}$$

なる事象を定義する. ここで, $\varphi_n(i)$ を入力としたときの通信路の出力を $\tilde{Y}^n(i)$ とおく. すると, メッセージ i が正しく復号される事象は, $E_i(\tilde{Y}^n(i)) \cap \left(\bigcap_{i' \neq i} E_{i'}^c(\tilde{Y}^n(i)) \right)$ と表すこと

ができる。したがって、メッセージ i を送信したときの復号誤り確率の符号に関する期待値 $\bar{\varepsilon}_n(i)$ は、

$$\begin{aligned}
\bar{\varepsilon}_n(i) &= \Pr \left\{ E_i^c(\tilde{Y}^n(i)) \cup \left(\bigcup_{i' \neq i} E_{i'}(\tilde{Y}^n(i)) \right) \right\} \\
&\leq \Pr \{ E_i^c(\tilde{Y}^n(i)) \} + \Pr \left\{ \bigcup_{i' \neq i} E_{i'}(\tilde{Y}^n(i)) \right\} \\
&\leq \Pr \{ E_i^c(\tilde{Y}^n(i)) \} + \sum_{i' \neq i} \Pr \{ E_{i'}(\tilde{Y}^n(i)) \} \\
&= \Pr \{ (\varphi_n(i), \tilde{Y}^n(i)) \notin \mathcal{T}_n \} + \sum_{i' \neq i} \Pr \{ E_{i'}(\tilde{Y}^n(i)) \} \tag{102}
\end{aligned}$$

ここで、 $\varphi_n(i)$ と $\tilde{X}^n$ は、同じ分布であり、かつ、 $\tilde{Y}^n(i)$ は、 $\varphi_n(i)$ を入力したときの通信路の出力であるので、 $\tilde{X}^n$ を入力したときの通信路の出力を $\tilde{Y}^n$ と表せば、

$$\Pr \{ (\varphi_n(i), \tilde{Y}^n(i)) \notin \mathcal{T}_n \} = \Pr \{ (\tilde{X}^n, \tilde{Y}^n) \notin \mathcal{T}_n \}. \tag{103}$$

よって、

$$\bar{\varepsilon}_n(i) \leq \Pr \{ (\tilde{X}^n, \tilde{Y}^n) \notin \mathcal{T}_n \} + \sum_{i' \neq i} \Pr \{ E_{i'}(\tilde{Y}^n(i)) \} \tag{104}$$

ここで、

$$\begin{aligned}
P_{Y^n}(\mathbf{y}) &= \sum_{\mathbf{x}} P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{x}} P_{X^n}(\mathbf{x}) W^n(\mathbf{y}|\mathbf{x}) \\
&\geq \sum_{\mathbf{x} \in \mathcal{A}_n} P_{X^n}(\mathbf{x}) W^n(\mathbf{y}|\mathbf{x}) \\
&= \kappa_n \sum_{\mathbf{x} \in \mathcal{A}_n} P_{\tilde{X}^n}(\mathbf{x}) W^n(\mathbf{y}|\mathbf{x}) \\
&= \kappa_n \sum_{\mathbf{x} \in \mathcal{A}_n} P_{\tilde{X}^n \tilde{Y}^n}(\mathbf{x}, \mathbf{y}) \\
&= \kappa_n P_{\tilde{Y}^n}(\mathbf{y}). \tag{105}
\end{aligned}$$

よって、

$$\begin{aligned}
\frac{1}{n} \log \frac{1}{P_{Y^n}(\tilde{Y}^n)} &\leq \frac{1}{n} \log \frac{1}{\kappa_n P_{\tilde{Y}^n}(\tilde{Y}^n)} \\
&= \frac{1}{n} \log \frac{1}{P_{\tilde{Y}^n}(\tilde{Y}^n)} + \frac{1}{n} \log \frac{1}{\kappa_n}. \tag{106}
\end{aligned}$$

よって,

$$\frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{Y^n}(\tilde{Y}^n)} \leq \frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{\tilde{Y}^n}(\tilde{Y}^n)} + \frac{1}{n} \log \frac{1}{\kappa_n}. \quad (107)$$

よって, 実数 a に対して,

$$\frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{\tilde{Y}^n}(\tilde{Y}^n)} + \frac{1}{n} \log \frac{1}{\kappa_n} < a \quad (108)$$

ならば,

$$\frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{Y^n}(\tilde{Y}^n)} < a. \quad (109)$$

よって,

$$\Pr \left\{ \frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{Y^n}(\tilde{Y}^n)} < a \right\} \geq \Pr \left\{ \frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{Y^n}(\tilde{Y}^n)} < a - \frac{1}{n} \log \frac{1}{\kappa_n} \right\}. \quad (110)$$

一方, 任意の $\mathcal{S} \subset \mathcal{X}^n \times \mathcal{Y}^n$ に対して,

$$\begin{aligned} \Pr\{X^n Y^n \in \mathcal{S}\} &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{S}} P_{X^n}(\mathbf{x}) W^n(\mathbf{y} | \mathbf{x}) \\ &\geq \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{S}: \mathbf{x} \in \mathcal{A}_n} P_{X^n}(\mathbf{x}) W^n(\mathbf{y} | \mathbf{x}) \\ &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{S}: \mathbf{x} \in \mathcal{A}_n} \kappa_n P_{\tilde{X}^n}(\mathbf{x}) W^n(\mathbf{y} | \mathbf{x}) \\ &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{S}} \kappa_n P_{\tilde{X}^n}(\mathbf{x}) W^n(\mathbf{y} | \mathbf{x}) \\ &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{S}} \kappa_n P_{\tilde{X}^n \tilde{Y}^n}(\mathbf{x}, \mathbf{y}) \\ &= \kappa_n \Pr\{\tilde{X}^n \tilde{Y}^n \in \mathcal{S}\} \end{aligned} \quad (111)$$

であるので, n が十分大きければ,

$$\begin{aligned} \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n | X^n)}{P_{Y^n}(Y^n)} < a \right\} &\geq \kappa_n \Pr \left\{ \frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{Y^n}(\tilde{Y}^n)} < a \right\} \\ &\geq \kappa_n \Pr \left\{ \frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{\tilde{Y}^n}(\tilde{Y}^n)} < a - \frac{1}{n} \log \frac{1}{\kappa_n} \right\} \\ &\geq \kappa_n \Pr \left\{ \frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{\tilde{Y}^n}(\tilde{Y}^n)} \leq a - \gamma \right\}. \end{aligned} \quad (112)$$

ここで,

$$a = I(X; Y) - \gamma \quad (113)$$

とおくと,

$$\begin{aligned} \Pr\{\tilde{X}^n \tilde{Y}^n \notin \mathcal{T}_n\} &= \Pr\left\{\frac{1}{n} \log \frac{W^n(\tilde{Y}^n | \tilde{X}^n)}{P_{\tilde{Y}^n}(\tilde{Y}^n)} \leq I(X; Y) - 2\gamma\right\} \\ &\leq \frac{1}{\kappa_n} \Pr\left\{\frac{1}{n} \log \frac{W^n(Y^n | X^n)}{P_{Y^n}(Y^n)} < I(X; Y) - \gamma\right\} \\ &\rightarrow 0 \quad (n \rightarrow \infty). \end{aligned} \quad (114)$$

一方,

$$\begin{aligned} \Pr\{E_{i'}(\tilde{Y}^n(i))\} &= \Pr\{(\varphi_n(i'), \tilde{Y}^n(i)) \in \mathcal{T}_n\} \\ &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n} \Pr\{\varphi_n(i') = \mathbf{x}\} \Pr\{\tilde{Y}^n(i) = \mathbf{y} | \varphi_n(i') = \mathbf{x}\} \\ &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n} P_{\tilde{X}^n}(\mathbf{x}) P_{\tilde{Y}^n}(\mathbf{y}) \\ &\leq \frac{1}{\kappa_n} \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n} P_{\tilde{X}^n}(\mathbf{x}) P_{\tilde{Y}^n}(\mathbf{y}) \\ &< \frac{1}{\kappa_n} \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n} P_{\tilde{X}^n}(\mathbf{x}) W^n(\mathbf{y} | \mathbf{x}) e^{-n(I(X; Y) - 2\gamma)} \\ &\leq \frac{1}{\kappa_n} e^{-n(I(X; Y) - 2\gamma)}. \end{aligned} \quad (115)$$

よって,

$$\begin{aligned} \sum_{i' \neq i} \Pr\{E_{i'}(\tilde{Y}^n(i))\} &\leq \frac{1}{\kappa_n} \sum_{i' \neq i} e^{-n(I(X; Y) - 2\gamma)} \\ &\leq M_n e^{-n(I(X; Y) - 2\gamma)} \\ &= e^{-n(I(X; Y) - 3\gamma)} e^{-n(I(X; Y) - 2\gamma)} \\ &= e^{-n\gamma}. \end{aligned} \quad (116)$$

よって,

$$\begin{aligned} \bar{\varepsilon}_n(i) &\leq \Pr\{(\tilde{X}^n, \tilde{Y}^n) \notin \mathcal{T}_n\} + \sum_{i' \neq i} \Pr\{E_{i'}(\tilde{Y}^n(i))\} \\ &\leq \Pr\{(\tilde{X}^n, \tilde{Y}^n) \notin \mathcal{T}_n\} + \frac{1}{\kappa_n} e^{-n\gamma}. \end{aligned} \quad (117)$$

式 (117) は, i によらないことに注意すれば, 復号誤り確率のすべての符号にわたる平均 $\bar{\varepsilon}_n$ は,

$$\begin{aligned}\bar{\varepsilon}_n &= \frac{1}{M_n} \sum_{i=1}^{M_n} \bar{\varepsilon}_n(i) \\ &\leq \frac{1}{M_n} \sum_{i=1}^{M_n} \left(\Pr\{(\tilde{X}^n, \tilde{Y}^n) \notin \mathcal{T}_n\} + \frac{1}{\kappa_n} e^{-n\gamma} \right) \\ &= \Pr\{(\tilde{X}^n, \tilde{Y}^n) \notin \mathcal{T}_n\} + \frac{1}{\kappa_n} e^{-n\gamma}.\end{aligned}\tag{118}$$

よって,

$$\lim_{n \rightarrow \infty} \bar{\varepsilon}_n = 0.\tag{119}$$

すべての n に対して, 復号誤り確率 ε_n が

$$\varepsilon_n \leq \bar{\varepsilon}_n\tag{120}$$

を満たす符号が存在するので,

$$\lim_{n \rightarrow \infty} \varepsilon_n = 0.\tag{121}$$

よって, レート R は Γ -達成可能である. すなわち,

$$\begin{aligned}C(\Gamma) &\geq R \\ &= I(X; Y) - 3\gamma\end{aligned}\tag{122}$$

X の任意性により,

$$C(\Gamma) \geq \sup_{X: E[c(X)] \leq \Gamma - \tilde{\gamma}} I(X; Y).\tag{123}$$

ここで, 右辺は, 補題 2 より, $\tilde{\gamma}$ に関して連続であり, かつ, $\tilde{\gamma} > 0$ は任意であったので,

$$C(\Gamma) \geq \sup_{X: E[c(X)] \leq \Gamma} I(X; Y)\tag{124}$$

となる. $\square$

A.2 逆定理の証明

証明 レート R が Γ -達成可能であるとする. すなわち, 式 (9),(10),(11) を満たす符号の列が存在する. このとき,

$$\varepsilon_n = \Pr\{X^n \neq \hat{X}^n\}\tag{125}$$

とすれば,

$$\begin{aligned}
\log M_n &= H(X^n) \\
&= H(X^n) - H(X^n|\hat{X}^n) + H(X^n|\hat{X}^n) \\
&= I(X^n; \hat{X}^n) + H(X^n|\hat{X}^n) \\
&\leq I(X^n; Y^n) + H(X^n|\hat{X}^n) \\
&\leq I(X^n; Y^n) + \varepsilon_n \log M_n + h(\varepsilon_n) \\
&\leq \frac{1}{1 - \varepsilon_n} (I(X^n; Y^n) + h(\varepsilon_n)).
\end{aligned} \tag{126}$$

よって,

$$\begin{aligned}
R &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \\
&\leq \liminf_{n \rightarrow \infty} \frac{1}{n} (I(X^n; Y^n) + h(\varepsilon_n)) \\
&= \liminf_{n \rightarrow \infty} I(X^n; Y^n).
\end{aligned} \tag{127}$$

ここで,

$$\begin{aligned}
I(X^n; Y^n) &= H(Y^n) - H(Y^n|X^n) \\
&\leq \sum_{i=1}^n H(Y_i) - H(Y^n|X^n) \\
&= \sum_{i=1}^n H(Y_i) - \sum_{i=1}^n H(Y_i|X^n Y^{i-1}) \\
&= \sum_{i=1}^n H(Y_i) - \sum_{i=1}^n H(Y_i|X_i) \\
&= \sum_{i=1}^n I(X_i; Y_i).
\end{aligned} \tag{128}$$

ここで,

$$P_Q(i) = \frac{1}{n}, \quad i = 1, 2, \dots, n \tag{129}$$

なる確率変数 Q を導入する． X^n とは独立である． よって，

$$\begin{aligned}
E[c(X_Q)] &= \sum_x c(x) P_{X_Q}(x) \\
&= \sum_x c(x) \sum_{i=1}^n P_{X_Q Q}(x, i) \\
&= \sum_x c(x) \sum_{i=1}^n P_{X_i Q}(x, i) \\
&= \sum_x c(x) \sum_{i=1}^n P_Q(i) P_{X_i}(x) \\
&= \frac{1}{n} \sum_{i=1}^n E[c(X_i)] \\
&= E \left[\frac{1}{n} c(X^n) \right] \\
&\leq \Gamma.
\end{aligned} \tag{130}$$

よって，

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n I(X_i; Y_i) &= \sum_{i=1}^n P_Q(i) I(X_i; Y_i) \\
&= \sum_{i=1}^n P_Q(i) (H(Y_i) - H(Y_i|X_i)) \\
&= \sum_{i=1}^n P_Q(i) H(Y_Q|Q=i) - \sum_{i=1}^n P_Q(i) H(Y_i|X_i) \\
&= H(Y_Q|Q) - \sum_{i=1}^n P_Q(i) \sum_x P_{X_i}(x) H(Y_i|X_i=x) \\
&= H(Y_Q|Q) - \sum_{i=1}^n P_Q(i) \sum_x P_{X_i}(x) H(Y_Q|X_Q=x) \\
&= H(Y_Q|Q) - \sum_x \sum_{i=1}^n P_Q(i) P_{X_i}(x) H(Y_Q|X_Q=x) \\
&= H(Y_Q|Q) - \sum_x P_{X_Q}(x) H(Y_Q|X_Q=x) \\
&= H(Y_Q|Q) - H(Y_Q|X_Q) \\
&\leq H(Y_Q) - H(Y_Q|X_Q) \\
&= I(X_Q; Y_Q).
\end{aligned} \tag{131}$$

よって,

$$\begin{aligned}
R &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^n) \\
&\leq \liminf_{n \rightarrow \infty} I(X_Q; Y_Q) \\
&\leq \liminf_{n \rightarrow \infty} \sup_{X: E[c(X)] \leq \Gamma} I(X; Y) \\
&= \sup_{X: E[c(X)] \leq \Gamma} I(X; Y).
\end{aligned} \tag{132}$$

すなわち, R が Γ -達成可能ならば,

$$R \leq \sup_{X: E[c(X)] \leq \Gamma} I(X; Y). \tag{133}$$

よって,

$$C(\Gamma) \leq \sup_{X: E[c(X)] \leq \Gamma} I(X; Y) \tag{134}$$

となる. $\square$

付録 B 定理 2 の証明

$$C = \sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} \tag{135}$$

を順定理および逆定理に分けて証明する. その後,

$$C = \sup_X \frac{I(X; Y)}{E[c(X)]} \tag{136}$$

を証明する.

B.1 順定理の証明

証明 任意の $\Gamma > 0$ に対して, $\frac{C(\Gamma)}{\Gamma}$ がコストあたり達成可能であることを示す.

任意の $\gamma > 0$ を考える. すると, $C(\Gamma) - \frac{\gamma\Gamma}{2}$ は, Γ -達成可能であるので,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \geq C(\Gamma) - \frac{\gamma\Gamma}{2} \tag{137}$$

なる固定長符号の列 $\{\varphi_n\}_{n=1}^{\infty}$ が存在する. よって, 十分大きなすべての n で

$$\frac{1}{n} \log M_n \geq C(\Gamma) - \gamma\Gamma. \tag{138}$$

十分大きな ν を考える．すると，

$$n\Gamma \leq \nu < n\Gamma + \Gamma \quad (139)$$

なる十分大きな n が一意に定まる．

ここで， ν に対応する符号 φ_ν を

$$\varphi_\nu = \varphi_n \quad (140)$$

とする．この φ_ν は，

$$\begin{aligned} c(\varphi_\nu(m)) &= c(\varphi_n(m)) \\ &\leq n\Gamma \\ &\leq \nu \end{aligned} \quad (141)$$

を満たす．また，

$$\begin{aligned} \frac{1}{\nu} \log M_\nu &> \frac{n}{\nu} (C(\Gamma) - \gamma\Gamma) \\ &> \frac{n}{n\Gamma + \Gamma} (C(\Gamma) - \gamma\Gamma) \\ &= \frac{n}{n+1} \left(\frac{C(\Gamma)}{\Gamma} - \gamma \right) \\ &> \frac{C(\Gamma)}{\Gamma} - 2\gamma \end{aligned} \quad (142)$$

であるので，

$$\liminf_{n \rightarrow \infty} \frac{1}{\nu} \log M_\nu \geq \frac{C(\Gamma)}{\Gamma} - 2\gamma \quad (143)$$

また，自明に

$$\begin{aligned} \lim_{n \rightarrow \infty} \varepsilon_\nu &= \lim_{n \rightarrow \infty} \varepsilon_n \\ &= 0 \end{aligned} \quad (144)$$

が成り立つ．

よって， $\frac{C(\Gamma)}{\Gamma} - 2\gamma$ は，コストあたり達成可能であるので，

$$C \geq \frac{C(\Gamma)}{\Gamma} - 2\gamma \quad (145)$$

$\gamma > 0$ は任意であったので，

$$C \geq \frac{C(\Gamma)}{\Gamma} \quad (146)$$

よって，

$$C \geq \sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} \quad (147)$$

となる．□

B.2 逆定理の証明

証明 R がコストあたり達成可能であると仮定する．すなわち，式 (18),(19),(20) を満たす符号の列 $\{\varphi_\nu\}_{\nu>0}$ が存在する．このとき，ファノの不等式より， φ_ν は，

$$\log M_\nu \leq \frac{1}{1-\varepsilon_\nu} (I(X^n; Y^n) + h(\varepsilon_\nu)) \quad (148)$$

を満たす．ただし， n は φ_ν の符号語長で，

$$X^n = \varphi_\nu(U_{M_\nu}) \quad (149)$$

Y^n は X^n を入力とする通信路の出力である． X^n は，

$$E[c(X^n)] \leq \nu \quad (150)$$

を満たすので，

$$\begin{aligned} \frac{1}{\nu} \log M_\nu &\leq \frac{1}{1-\varepsilon_\nu} \left(\frac{1}{\nu} I(X^n; Y^n) + \frac{1}{\nu} h(\varepsilon_\nu) \right) \\ &\leq \frac{1}{1-\varepsilon_\nu} \left(\frac{1}{\nu} \sup_{X^n: E[c(X^n)] \leq \nu} I(X^n; Y^n) + \frac{1}{\nu} h(\varepsilon_\nu) \right) \\ &= \frac{1}{1-\varepsilon_\nu} \left(\frac{n}{\nu} \sup_{X^n: E[\frac{1}{n} c(X^n)] \leq \frac{\nu}{n}} \frac{1}{n} I(X^n; Y^n) + \frac{1}{\nu} h(\varepsilon_\nu) \right) \end{aligned} \quad (151)$$

ここで，

$$P_Q(i) = \frac{1}{n}, \quad i = 1, 2, \dots, n \quad (152)$$

なる確率変数 Q を導入する． X_Q と同じ分布で互いに独立な確率変数 $\tilde{X}_i$ ($i = 1, 2, \dots, n$) を定めて，

$$\tilde{X}^n = \tilde{X}_1 \tilde{X}_2 \cdots \tilde{X}_n \quad (153)$$

とする。すると、

$$\begin{aligned}
E[c(X_Q)] &= \sum_{i=1}^n P_Q(i) E[c(X_Q)|Q=i] \\
&= \frac{1}{n} \sum_{i=1}^n E[c(X_i)] \\
&= \frac{1}{n} E \left[\sum_{i=1}^n c(X_i) \right] \\
&= \frac{1}{n} E[c(X^n)] \\
&\leq \frac{\nu}{n}
\end{aligned} \tag{154}$$

であるので、

$$\begin{aligned}
E[c(\tilde{X}^n)] &= \sum_{i=1}^n E[c(X_Q)] \\
&\leq \nu
\end{aligned} \tag{155}$$

となる。さらに、

$$\begin{aligned}
\frac{1}{n} I(X^n; Y^n) &\leq \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i) \\
&\leq I(X_Q; Y_Q)
\end{aligned} \tag{156}$$

および

$$E[c(X_Q)] \leq \frac{\nu}{n} \tag{157}$$

が成り立つので、

$$I(X_Q; Y_Q) \leq \sup_{X: E[c(X)] \leq \frac{\nu}{n}} I(X; Y) \tag{158}$$

となり、よって、

$$\begin{aligned}
\frac{1}{\nu} \log M_\nu &\leq \frac{1}{1-\varepsilon_\nu} \left(\frac{n}{\nu} \sup_{X: E[c(X)] \leq \frac{\nu}{n}} I(X; Y) + \frac{1}{\nu} h(\varepsilon_\nu) \right) \\
&\leq \frac{1}{1-\varepsilon_\nu} \left(\sup_{\Gamma > 0} \frac{1}{\Gamma} \sup_{X: E[c(X)] \leq \Gamma} I(X; Y) + \frac{1}{\nu} h(\varepsilon_\nu) \right) \\
&\leq \frac{1}{1-\varepsilon_\nu} \left(\sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} + \frac{1}{\nu} h(\varepsilon_\nu) \right)
\end{aligned} \tag{159}$$

となる。したがって,

$$\begin{aligned} R &\leq \liminf_{\nu \rightarrow \infty} \frac{1}{\nu} \log M_\nu \\ &\leq \sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} \end{aligned} \tag{160}$$

となり,

$$\begin{aligned} \sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} &\geq \sup\{R \mid R \text{ がコストあたり達成可能} \} \\ &= C \end{aligned} \tag{161}$$

と得る。□

B.3 式 (136) の証明

証明 式 (135) より,

$$\begin{aligned} \sup_{\Gamma > 0} \frac{C(\Gamma)}{\Gamma} &= \sup_{\Gamma > 0} \frac{1}{\Gamma} \sup_{X: E[c(X)] \leq \Gamma} I(X; Y) \\ &= \sup_{\Gamma > 0} \sup_{X: E[c(X)] \leq \Gamma} \frac{I(X; Y)}{\Gamma} \\ &= \sup_X \sup_{\Gamma \geq E[c(X)]} \frac{I(X; Y)}{\Gamma} \\ &= \sup_X \frac{I(X; Y)}{E[c(X)]} \end{aligned} \tag{162}$$

となる。□